ISPyB Developer's meeting

April 29 2019

DRAFT

Participants:

ALBA: Daniel Sanchez

Diamond: Karl Levik, Neil Smith, NN,

EMBL-HH: Ivars Karpics

ESRF: Olof Svensson, Solange Delageniere, Alejandro de Maria, Maxim

Global Phasing: Clemens Vonrhein, Peter Keller, Rasmus Fogh

SOLEIL: Idrissou Chado, Tatiana Isabet

1. Setting the agenda.

The agenda was agreed

2. Minutes and matters arising

The minutes of the last developers' meeting (at the Lund conference) were agreed.

3. Status reports

SOLEIL: Working on setting up reprocessing

ALBA: Should have caught up with current version now

EMBL-HH: Nothing to report

Global Phasing: Co-organised Data Metrics meeting in Grenoble, with ESRF (see point 4)

ESRF: Also working on reprocessing and proposes collaboration and ideas sharing with SOLEIL on this point. Moved to new react framework.

Diamond: New system for displaying diffraction images, working purely from HDF5 files. The Scipion launcher for EM will be ready RSN. Diamond has been looking at Vuera, a program to reuse components between Vue and React that should allow sharing of components. See e.g.https://x-team.com/blog/react-vue-component-integration/

4. Report of the Data Metrics meeting in Grenoble

The Metrics was co-organised by Global Phasing and the ESRF. The meeting brought together theoreticians and developers, synchrotron people, users' representatives and a PDB representative, and discussed the needs for improved practices and data metrics. Presentations and a discussion summary are on a draft web site, which will be opened to participants RSN, as soon as a first round of consultation on the results has finished. A number of ISPyB-related points were raised at the meeting:

- Reducing the number of parallel autoprocessing calculations made and presented, to make it easier for users to decide what to look at.
- More user control over reprocessing
- Better diagnostics from processing programs, especially when they fail.
- The need for better data harvesting - whether through image headers., ISPyB, or mmCIF files.
- User requests for more standardisation on access, viewing, and download from different synchrotrons, and for simpler downloading procedures.

Several participants stressed the need for more resources to ISPyB development.

In the discussion it was confirmed that the open source version of MRFANA was in progress. A first step will be to improve documentation, so that MRFANA can be used by third parties as-is, pending a code clean-up to finalise the open source version.

5. Review of issues

The discussion of the specific issues will continue on Github, and people are requested to contribute.

#43 New column for AutoProcProgram Attachment. It was agreed that the proposed 'primary' Boolean should be replaced by a ranking integer, as proposed by CV. It was discussed, but rejected, to add a uniqueness constraint to the new column.

#44 BLSession.archived new Boolean column The purpose of the column would be to distinguish cases of missing files from cases where files had been archived and removed from the server according to normal procedures. The scope would be only for displayed results data, not for raw data. It would be relevant only for Diamond, as ESRF uses only small, thumbprint imagers and never removes any of them. It was discussed, but not decided, to combine the flag with a status enumeration.

 # 42 Schema fixing: Rename Protein, Crystal and DiffractionPlan. The advantage of cleaning up code and naming, removing technical debt, was generally agreed. Specific measures were more controversial. The Steering Committee had suggested limiting development to cater for the techniques MX, EM, and SAXS, but this still would benefit for generalisation and harmonisation. AdM objected that the Crystal table was used in a number of different components, and a renaming would require considerable work for negligible practical advantage. As regards the Protein table, it was agreed that renaming c/should be combined with a reworking to support more complex needs, such as complexes. It was agreed that 'Macromolecule' would be a better name for the modified

table than 'Component', and that anything that allowed merging of the current Protein and Biosax tables would be a 'good thing'. Diamond and ESRF hope to produce a joint proposal before the next developers' meeting.

Other desirable changes were discussed. One proposal was to rethink (rather than merely patch) the Container table. Another was to reanalyse the Screening tables, where there are currently 8-10 tables but only 3-4 fields were used. Yet another to merge the tables for data collection, energy scans, and XRF. Another point of interest would be to gather use cases for serial crystallography.

4. Any Other Business

Considering that any major changes would have to be bedded in before the ESRF restarted operations, it was proposed to organise a refactoring face-to-face meeting of the developers to 'make changes more likely' (OS). Paris was selected as a good, centrally located location, and SOLEIL agreed to organise meeting rooms, accommodation etc. The meeting should be in September and run over two days, one night, to allow travel on the morning of the first and afternoon of the second day. RF has set up a Doodle poll for suitable dates in September (https://doodle.com/poll/tq8tiay8i23emet7). NOTE that the poll asks for the **first** day of a two-day meeting.

It was emphasised that the success of the face-to-face meeting would hang on thorough preparation of the issues to discuss, and doing homework beforehand.

Next Meeting

The next meeting should be in the first week of June; RF has set up a Doodle poll (https://doodle.com/poll/2i82d379smc2bh3p).

Rasmus Fogh, Global Phasing   30/04/19